



University of Pennsylvania
ScholarlyCommons

Center for Human Modeling and Simulation

Department of Computer & Information Science

July 2002

FacEMOTE: Qualitative Parametric Modifiers for Facial Animations

Norman I. Badler

University of Pennsylvania, badler@seas.upenn.edu

Meeran Byun

University of Pennsylvania

Follow this and additional works at: <http://repository.upenn.edu/hms>

Recommended Citation

Badler, N. I., & Byun, M. (2002). FacEMOTE: Qualitative Parametric Modifiers for Facial Animations. Retrieved from <http://repository.upenn.edu/hms/52>

Postprint version. Copyright ACM, 2002. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in *Symposium on Computer Animation*, pages 65-71.

Publisher URL: <http://doi.acm.org/10.1145/S45261.S45272>

This paper is posted at ScholarlyCommons. <http://repository.upenn.edu/hms/52>

For more information, please contact libraryrepository@pobox.upenn.edu.

FacEMOTE: Qualitative Parametric Modifiers for Facial Animations

Abstract

We propose a control mechanism for facial expressions by applying a few carefully chosen parametric modifications to preexisting expression data streams. This approach applies to any facial animation resource expressed in the general MPEG-4 form, whether taken from a library of preset facial expressions, captured from live performance, or entirely manually created. The MPEG-4 Facial Animation Parameters (FAPs) represent a facial expression as a set of parameterized muscle actions, given as intensity of individual muscle movements over time. Our system varies expressions by changing the intensities and scope of sets of MPEG-4 FAPs. It creates variations in “expressiveness” across the face model rather than simply scale, interpolate, or blend facial mesh node positions. The parameters are adapted from the Effort parameters of Laban Movement Analysis (LMA); we developed a mapping from their values onto sets of FAPs. The FacEMOTE parameters thus perturb a base expression to create a wide range of expressions. Such an approach could allow real-time face animations to change underlying speech or facial expression shapes dynamically according to current agent affect or user interaction needs.

Keywords

Facial animation, Animation systems, MPEG

Comments

Postprint version. Copyright ACM, 2002. This is the author's version of the work. It is posted here by permission of ACM for your personal use. Not for redistribution. The definitive version was published in *Symposium on Computer Animation*, pages 65-71.

Publisher URL: <http://doi.acm.org/10.1145/545261.545272>

FacEMOTE:
Qualitative Parametric Modifiers for Facial Animations

Meeran Byun

Computer and Information Science
University of Pennsylvania*

Norman I. Badler

Abstract

We propose a control mechanism for facial expressions by applying a few carefully chosen parametric modifications to pre-existing expression data streams. This approach applies to any facial animation resource expressed in the general MPEG-4 form, whether taken from a library of preset facial expressions, captured from live performance, or entirely manually created. The MPEG-4 Facial Animation Parameters (FAPs) represent facial expression as a set of parameterized muscle actions, given as intensity of individual muscle movements over time. Our system varies expressions by changing the intensities and scope of sets of MPEG-4 FAPs. It creates variations in “expressiveness” across the face model rather than simply scale, interpolate, or blend facial mesh node positions. The parameters are adapted from the Effort parameters of Laban Movement Analysis (LMA); we developed a mapping from their values onto sets of FAPs. The FacEMOTE parameters thus perturb a base expression to create a wide range of expressions. Such an approach could allow real-time face animations to change underlying speech or facial expression shapes dynamically according to current agent affect or user interaction needs.

Keywords: Facial animation, Animation systems, MPEG

1 INTRODUCTION

Facial animation has been an area of active research almost since the beginning of 3D graphics. Various approaches have been proposed to generate realistic facial animations since the seminal work by Parke[1974]. Facial animation research includes generating face geometry models, finding mechanisms to move surface points realistically on the face, applying lip synchronization to speech or other vocalizations, and animating other facial attributes such as jaw rotation, eye movement, and vascular changes. A number of techniques have been developed for creating photorealistic face models from 2D images [Blanz and Vetter 1999; Lee and Thalmann 2000; Pighin et al. 1998]. These techniques enable custom-building a 3D texture-mapped face model from one or more images with high degree of realism. Lip-synching software allows simple creation of animated lips to a speech signal REF??[lipsinc.com]. Parametric animation techniques manipulate facial points or skin regions by geometric interpolation, muscle models, or deformable surfaces [Kalra et al. 1992; Lee et al. 1995; Parke 1974; Pighin et al. 1998; Platt and Badler 1981; Terzopoulos and Waters 1991; Waters 1987]. Hybrid approaches retarget or blend a given expression from one source face onto another of similar or differing mesh topology.

Face animation research has mainly attempted to incorporate a few general principles into face animation systems:

- The animations generated by the system should be realistic or believable: for example, emotions should be “readable” from the face; eyelids should blink appropriately; and lips, teeth, and tongue should be modeled and properly animated.
- A wide range of expressions must be created: emotions as well as puckering, chewing, and blushing.
- Animation control parameters should be intuitive and easy to use: the best example is using the speech signal itself to control changing lip shapes.
- Control parameters should be consistent and applicable across different face models: adapting animation data for a different face model should require as little manual intervention as possible.

Performance-driven methods use electromechanical devices to capture face movement directly from marked surface points or vision techniques to extract facial motion in real-time or from video. These methods are useful for generating realistic animations for a specific target model or for building a library of expressions. However, they need elaborate initial set up to generate animation data and retargeting the data for a new face model requires considerable work. Another group of approaches includes muscle simulation methods that approximate the effects of muscle actions on the facial skin using various techniques. These methods generate facial animation by manipulating the face geometry mesh with muscle action simulation and shape transformation. Physics-based models simulate mechanical properties of skin tissues and their underlining facial muscles. Spring meshes [Lee et al. 1995; Platt and Badler 1981; Terzopoulos and Waters 1991] or vectors [Waters 1987] are used to simulate muscle action. Each muscle is possibly activated and the effect of muscle activations on the elastic skin is computed for animation synthesis. Mathematical approximation models include free form deformation [Kalra et al. 1992] to approximate changes in skin surface shape due to muscle movements. Geometric interpolation methods synthesize key poses from various sources and interpolate between them [Parke 1974; Pighin et al. 1998]. The methods discussed can generate facial animations by controlling individual muscles, animation parameters, or control points, which manipulate facial regions locally. Thus sculpting facial animation or fine-tuning existing facial animation data to fit individual needs using these methods produces highly realistic results but requires considerable manual work by animators.

Parameterization techniques have been an area of active research since Parke [Parke 1974]. Two parameter sets, Facial Action Coding System (FACS) [Ekman and Friesen 1977] and more recently Facial Animation Parameters (FAP) [MPEG-4 SNHC 1998] are widely used as *de facto* standards for describing facial motions. FAPs provide a set of independent low level parameters that are designed to animate faces of different sizes and proportions. Since the units of animation parameters are defined in terms of the distances between facial feature points of a given face model, applying the same FAP data on different face models produces a high quality result even with differences in facial morphology [Ostermann 1998]. However, the large number of parameters (66) makes it awkward to use FAPs as a direct (interactive) synthesis tool.

We propose a new method for facial animation synthesis in which a small set of high level parameters drive sets of underlying

*{byun,badler}@seas.upenn.edu

low level FAPs. Our goal is to simplify the process of facial animation synthesis by allowing high level parametric control over the whole face. The key is to identify a reasonable set of such high level parameters [Parke and Waters 1996, p. 190]. We adapt the EMOTE system [Chi et al. 2000], which was developed for arm gestures and postures, for the face, and call the synthesis FacEMOTE. We use FAP streams as the building blocks for animation, run a set of filtering functions on the input FAP stream, and generate a new FAP stream as output. We try to capture “expressiveness” of facial movements in a set of filtering functions controlled by four high level EMOTE parameters. Thus users can generate subtle differences in facial animations by simply controlling those four parameters, creating a wide range of facial animations starting from a single animation sequence. This system makes the job of facial animation synthesis much easier by two main factors: reducing the number of parameters to a small number of intuitive parameters and applying the parameters over a sequence of time rather than on a single frame, thus producing animation sequences for playback in real time. Moreover, the original base expression in terms of FAPs is still available and modifiable; the FacEMOTE parameters place another layer of control on top of them.

Noh et al. [2001] proposed a method for retargeting facial expressions onto new models. This method maps existing expressions to new models regardless of morphology, e.g., whether human or animal, but doesn’t manipulate the expression data to create a new one. Thus, to create novel expressions, a library with a large number of expressions and an expression blending tool are still needed. Essa et al. [1997] improved on FACS, a standard parameterization system, by adding temporal patterning of muscle actuations and muscle coarticulation in the eye region observed from empirical data. But the coarticulation data is tied to the analysis using optical flow on a small set of expressions (e.g., smile or surprise) and can be used for general purpose synthesis to a limited degree. Perlin’s work on the Improv face [mrl.nyu.edu/~perlin] is the closest to our work in terms of high level modification and augmentation of existing animations. He uses carefully constructed noise functions using statistical modeling to animate movements, such as eye blinking, to add life-like quality to simple keyframe animation sequences. Simple but powerful numerical compositing operations algebraically combine differing animation sequences and allow simultaneous expressions. The main differences between this work and ours is that we carefully choose specific parameters based on movement observation models to add “expressive qualities” to the base expressions, our face models are inherently 3D, and we permit the modification of external sources of facial data streams.

We discuss the design motivation in Section 2, then what facial parameters or sets of parameters are affected by EMOTE parameters in Section 3. In Section 4 we discuss the mappings that relate EMOTE parameters to FAPs, followed by discussions on the results and conclusions.

2 Design Motivation

Among the various channels of human communication such as facial expressions, gestures, postures, vocalizations, and vascular expressions, facial expression is one in which a spectrum of many

different expressions are possible, sometimes at a speed human eyes can barely discern (micromomentary expression). In the framework of parameterizing facial animations, we would like to have the following minimum characteristics to capture the “expressiveness” of facial movements:

- An expression’s development over time at various speeds is integrated. For example, micromomentary expression happens when a person wants to hide his true feelings [Knapp 1992]. It is known that people can judge expressions more accurately when they are shown dynamic images than a single static image [Wehrle 2000].
- The relative strength of expressions can be controlled.
- Facial expression parameters may be extended to the rest of the body to keep consistency in other communication channels. Researchers from various fields agree that if consistency of expression among different channels is violated, the expression will look insincere or forced. It can have a damaging effect to the plausibility or believability of the animation [Frijda and Tcherkassof 1997; Isbister and Nass 2000]. Animators know that capturing expression in the face and throughout the body is an important factor in making characters come alive [Lasseter 1987; Thomas and Johnston 1981].

The EMOTE system [Chi et al. 2000], a parameterized system for gesture and body posture, inspired us to design a complementary framework for facial actions. The parameters for the EMOTE system satisfy the first two conditions. The third condition was more elusive, since little seems to have been published on any direct relationship between Laban Movement Analysis (LMA) and facial actions. Bartenieff [Bartenieff and Lewis1980] (a contemporary of Laban’s) states, however, that a preparative or a held attitude can be observed in the face and hands and it continues into the rest of the body. For example, a combination of LMA direct and bound qualities (explained in the next Section) can be observed in the stare of the eyes and also in the whole body; the shoulders are lifted and narrowed toward the neck, and the torso leans rigidly forward. Our extension of EMOTE to FacEMOTE does not benefit from the large body of empirical observations of decades of LMA professionals; rather we must rely on the principle of consistency of expression across the body’s communication channels to develop the requisite mappings from a similar set of LMA parameters applied as facial expression modifiers.

3 Movement Parameters

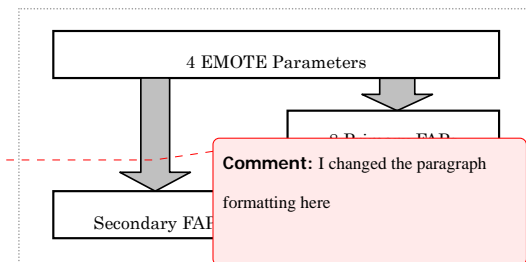


Figure 2. Parameter organization

Space	Indirect	Scanning the party floor. Rolling the eyes with no particular focus.
	Direct	Focusing on a ball player at the ball field. Squinting at the object an artist is drawing. Blowing out a candle.
Weight	Light	Whispering to a child to sleep. Lightly tickled into giggling. Whining in a muffled sound. Licking ice cream.
	Strong	Spelling out a word at a spelling bee. Snarling at an offender. Putting on a stern face when scolding a child.
Time	Sustained	Relaxed expression while daydreaming. Taking a deep breath. Yawning.
	Quick	Nervous fidgeting. Coughing. Clearing the throat. Sobbing of a child after a screaming fit.
Flow	Free	Crying of a baby when it is hungry. Bursting into uncontrollable laughter. Shouting in raging fury.
	Bound	Holding back tears. Chuckling instead of laughing loudly. Grimacing when touching a repulsive object.

Table 1. Effort elements

The EMOTE system is a computational interpretation of LMA, a method for describing qualitative changes in movements. The EMOTE system implementation of LMA generates arm and torso movements by varying four Effort and four Shape parameters. For example, a “careful” movement may be described by a combination of Light and Sustained Effort in arm movements and Retreating Shape displayed by the torso. Effort describes how one moves, or how one’s exertion is directed within functional or expressive behavior. Shape describes the form of the movement or how the body changes and moves through space (spreading or enclosing, rising or sinking, advancing or retreating) [Bartenieff and Davis 1972; Dell 1977]. For a human face, the Shape terms appear less descriptively useful, though they could clearly relate to head attitude with respect to the torso. Effort has more apparent utility for the deformable aspects of a face. We concentrate on only Effort parameters to describe facial movements.

3.1 Effort Parameters

Effort consists of four motion parameters: Space, Weight, Time, and Flow. Change in each parameter takes place in a range between two opposite extremes, or Effort elements. We use Effort parameters of values ranging from -1 to 1 to control expressive movements of the face. Table 1 lists the Effort parameters along with the Effort elements per parameter, and some examples of facial expressions or movements illustrating the Effort elements. The examples are compiled from literature on LMA and on facial expressions [Bartenieff and Lewis 1980; Faigin 1990; Peck 1990]. We chose some examples from other sources that can be clearly categorized as expressions displaying Effort elements to supplement material taken from the LMA literature.

3.2 MPEG-4 FAPs

A set of FAPs specifies an expression by each FAP indicating an atomic facial movement. A neutral expression is specified by a set of FAPs of zero values, and a FAP of non-zero value represents a

facial movement deforming a part of face from its relaxed state. For example, a sequence of facial actions acquired by increasing and then decreasing the value of ‘close_t_l_eyelid’ will make the model wink with its left eye. The values of FAPs are defined in Face Animation Parameter Units, or FAPUs – eye separation, eye-nose separation, mouth nose separation, and mouth width of the face model in its relaxed state – allowing independence of animation from facial morphology.

FAPs are arranged into 1 high level and 9 low level groups, each of the 9 low level groups representing a region on the face (1: visemes and expressions; 2: jaw, chin, inner lowerlip, cornerlips, midlip; 3: eyeballs, pupils, eyelids; 4: eyebrow; 5: cheeks; 6: tongue; 7: head rotation; 8: outer lip positions; 9: nose; 10: ears). A total of 26 parameters (16 + 10) from 2 groups (group 2 and 8) out of 66 low level parameters are designated on the mouth region for proper animation of talking heads.

We do not include FAPs for the regions where their movements are mainly functional or visible only at close range (tongue, nose, ears, pupils) and 2 high level parameters (visemes, expressions) in our system. Thus we have a total of 47 FAPs affected by Effort parameters.

3.3 Parameter Organization

It is a daunting task to map from a small number of high level parameters to each of 47 low level FAPs, also ensuring smooth transitions between regions deformed by each parameter. We simplify the task by observing that a set of mappings for a minimum number of parameters can be isolated. If the mappings in the minimum set are provided, then the rest of mappings may be obtained by blending a subset of mappings from the minimum set. Note that parameter organization is not used for compressing data by interpolating FAP values. Instead, it is used for interpolating functions that modulate an already existing animation given as a FAP stream, by exploiting dependency between parameters.

We categorize the 47 parameters into 2 groups: primary parameters representing the minimum set of independent parameters, and secondary parameters representing the rest which can be interpolated from the primary parameters. We selected 8 primary parameters by analyzing 38 seconds of motion captured FAP data, concentrating on the location of the major muscles of facial expressions [Williams et al. 1989]. In choosing the primary FAPs, priority is given to FAPs that are easy to track in the data acquisition process thus having a higher chance of accuracy. The 8 primary FAPs are:

1. ‘open_jaw’,
2. ‘lower_t_midlip’,
3. ‘raise_b_midlip’,
4. ‘stretch_cornerlip’,
5. ‘raise_l_cornerlip’,
6. ‘close_t_l_eyelid’,
7. ‘raise_l_i_eyebrow’,
8. ‘squeeze_l_eyebrow’.

We do not include parameters for eyeballs and head since they do not affect the rest of the parameters deforming the facial tissues. All primary parameters are chosen arbitrarily from the left side making use of symmetry. The mappings for the right side may trivially be calculated by mirroring them.

Color Plate 1a and 1b show examples of dependency among parameters. In Plate 1a, 'raise_b_lm' mirrors 'open_jaw'; the movement of 'stretch_l_cornerlip' closely reflects that of 'open_jaw' with small differences at several places; the movement of 'raise_b_midlip' is independent. In Plate 1b, 'puff_l_cheek' and 'lift_l_cheek' reflect 'raise_l_cornerlip' or 'raise_l_o_eyebrow' at different places. The dependencies between the 8 primary parameters and each of the secondary parameters are assigned as weights for a blending function and kept in a lookup table.

4 System Implementation

Our system generates new animation data using a simple offline preparative step and then modifying input FAP stream in three main steps. The preparative step generates a table of movement ranges once per model to avoid cartoon-like animation and artifacts (although we may want to extrapolate or exceed these for caricatures). FAPs should be applicable across face models, but in reality, models may not be designed for exaggerated facial movements, and artifacts may occur depending on how they are implemented.

The system uses the simple observation that facial expressions are generated by combinations of facial muscle movements, where muscle movements are either contractions or relaxations. A switch between muscle contraction and relaxation is marked by a local peak in the FAP stream. In varying facial expressions, generating a variety of muscle actions is possible by maintaining the overall structure of muscle actions (contracting or relaxing of muscles) but changing the path of the muscle trajectory. The overall goal of our system is to isolate and to define a meaningful subspace of such muscle trajectory space.

The first step analyzes an input FAP stream for the profile of the movement trajectory per parameter. We find local minima and maxima defining the profile. The profiles are used in the next two steps in reshaping the input FAP stream. User input of EMOTE parameters is used in this step to refine choices of local peaks by setting the threshold values between peaks. The second step modulates the primary FAPs. The EMOTE parameters are multiplied by each modulation function to adjust the strength of modulation. The third step modulates the rest of the FAPs with the functions obtained by blending functions used for the primary FAPs. Blending weights reflect the relative influence of each primary parameter on a particular secondary parameter. For example, 'left_l_cheek' has weights '0.2' for 'raise_l_i_eyebrow', '0.4' for 'raise_l_cornerlip', '0.4' for 'open_jaw' and zeroes for the rest of the primary parameters. The FAP stream generated in these steps then can be applied on a face model for rendering.

4.1 FAP Stream Analysis

In this step, the movement trajectory for each parameter is analyzed in preparation for the next two steps. In later steps, applying an EMOTE parameter to an input FAP stream will reshape the movement trajectory by changing the intensity and the duration of each FAP locally. For example, increasing the value of 'Quick' (positive value of Time parameter) will shorten the time taken from the onset to the decay of local actions, but the lengths of the FAP streams remain the same. Likewise, increasing the value of 'Sustained' will lengthen the time taken.

Two things are considered in this step: where the local peaks are and whether each of the local peaks should be included in the

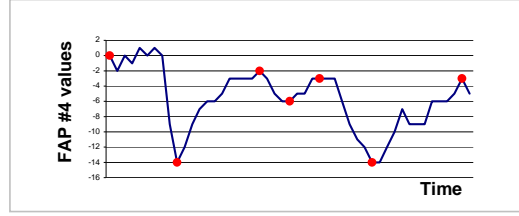


Figure 3. Local minima and maxima

movement profile or should be regarded as noise. While scanning the FAP stream for local peaks, locations of local minima and maxima exceeding both time and intensity thresholds are extracted. Local peaks whose values are smaller than thresholds are considered noise. The Time parameter has an influence on movement duration and the Flow parameter on movement intensity. Thus the values of Time and Flow parameters are considered in setting the thresholds. Thresholds for 'time' and 'intensity' are set relative to their initial values when Time and Flow parameters are both zero, defined by the following equations.

$$\varepsilon_i = \alpha \cdot \varepsilon_{i0} \quad (1a)$$

$$\varepsilon_i = \beta \cdot \varepsilon_{i0} \quad (1b)$$

Where ε_i , ε_i are the intensity and the time threshold values, and α , β are coefficients set by EMOTE parameters.

4.2 Primary FAP Modulation

As described in the previous Section, FacEMOTE works as a filter perturbing the value of each FAP for every frame as it runs through the FAP stream. A set of functions $\{F_1, \dots, F_8\}$ is computed for given values of EMOTE parameters as filters modulating each of the 8 primary parameters as follows.

$$F_i(s_i) = F_t^i \cdot F_f^i \cdot F_m^i(s_i) \quad (2)$$

Where F_f^i is a Flow function shaping i^{th} primary FAP stream;
 F_t^i is a Time function shaping i^{th} primary FAP stream;
 $F_m^i = a_i(s_i) + b_i$, s_i is the i^{th} primary FAP stream.

The coefficients for F_m^i , a_i and b_i are chosen from the coefficients of functions defined for each of the Space, Weight, Time, and Flow parameters. Their definitions (described as the coefficients a 's and b 's) at all of the Effort elements (two opposite extremes per parameter) are kept in a look-up table. In order to obtain F_m^i for given values of EMOTE parameters, function definitions for each of the EMOTE parameters are read from the table first. Then each of the 4 functions are linearly interpolated according to the value of the EMOTE parameters and the function that will maximize the modulation is chosen. For the neutral value of any EMOTE parameter, the values of a and b are set to 1 and 0 respectively, thus yielding the original animation.

The table for F_m^i of size 8 by 8 by 2 (for 8 primary parameters, 8 Effort elements, a and b) is built as follows: A set of frames are chosen at random. Then the facial expression when a single EMOTE parameter is at its extreme value is generated by using a facial expression editor using the heuristic rules given below. FAP values are collected from the generated expressions and linear

$$a_j = |e| \cdot \sum_{i=0}^8 w_i^j a_i \quad (4a)$$

$$b_j = |e| \cdot \sum_{i=0}^8 w_i^j b_i \quad (4b)$$

fitting is done on them to find the coefficients a and b .

- The function for the i^{th} primary parameter is defined independent of all other FAPs.
- For the Space parameter, the FAP groups 3 and 4 are the main focus.
- For the Weight parameter, the FAP group 2 is the main focus.
- For the Time parameter, Sustained value increases the intensities of FAPs while Quick value decreases the intensities.
- For the Flow parameter, Free value tends to open up facial features, thus increasing intensities, while Bound value tends to close in.

F_t^i and F_f^i are applied for non-zero Time and Flow values, shaping all FAPs locally. For the Free value of Flow parameter, extra movements (reflecting jaw movements) are added to eyebrows, corner-lips, and cheeks. The extra movements are fitted to a sine curve whose amplitude is proportional to that of the local peak.

For the Sustained value of the Time parameter, a Box filter of window size 4 is applied repeatedly. The number of repeats is determined by the value of the Time parameter. For the Quick value of the Time parameter, the following equation is used. Local minima and maxima extracted in the analysis step is used in this step to preserve the overall intensity. A threshold picked according to Time and Flow parameters have been used in the analysis to set the minimum time interval and the difference in intensity between neighboring local minima and maxima, thus affecting the overall shape of the FAP stream.

$$F_t^i(S_i) = \frac{M}{|M|} \cdot \left| \frac{S_i}{M} \right|^n \quad (3)$$

Where M is a local minimum or a maximum, and the value of n is determined by the Flow parameter.

4.3 Secondary FAP Modulation

Secondary FAPs are modulated by filtering functions of the form described by equation 2. F_m^i for a secondary function is computed by blending the functions for primary FAPs. The coefficients of the modulation function for each EMOTE parameter is assigned by the equations 4a and 4b. The blending weights for each primary FAP is looked up from the weight table. The weight table is generated by visually inspecting FAP graphs similar to Plate 1 and selecting primary parameters that the given secondary parameter is dependent on and assigning weight to each of them.

Where e is the value of an EMOTE parameter; and w_j^i is the weight of i^{th} primary parameter for j^{th} FAP.

F_t^i and F_f^i for the secondary functions are defined similar to those of the primary functions.

5 Results

We tested the system on a Pentium III PC by interactively setting the 4 EMOTE parameters and then playing the modified animation sequence. The FacEMOTE system automatically modifies the input FAP text stream in real time, outputting a FAP stream with varied "expression". The facial model is animated by simply playing the modified output FAP stream. The system was tested on a sample FAP file played at 25 frames per second. (The frame rate is the same for both the input and output file). Color Plate 2 shows snapshots of the original and modified animations over a time period of approximately 1.5 seconds. Table 2 gives a set of values for the EMOTE parameters for the examples shown in Plate 2.

	Space	Weight	Time	Flow	Qualitative description
(a)	0	0	0	0	Neutral
(b)	-1	0	-1	0	Indirect, sustained
(c)	0.5	1	0	0	Direct, strong
(d)	0	0	0	-1	Free

Table 2. EMOTE parameter values for Color plate 2

The changes in expressiveness in facial movements over time are captured in the output animation. Setting the Weight parameter to Strong, which involves jaw movements, results in more evident changes than ones involving strictly soft tissue movements.

6 Conclusion

We have introduced a parameterization for a facial animation system that creates a space of realistic facial animations by tuning a handful of intuitive control parameters. This is achieved by selecting Effort parameters, attributed to the expressiveness or qualitative aspects of movements, as the high level parameters controlling low level parameters that define local deformations. The implementation of the FacEMOTE system applies knowledge in facial physiology and heuristic methods in manipulating animation data. FacEMOTE takes advantage of an established parameterization, MPEG-4 FAP. It also generates many different shades of expression from a single base expression by adding subtle changes. FacEMOTE can be plugged into virtually any existing facial animation system that supports parameterization of soft tissue deformations to develop a comprehensive animation generation tool with real time playback capability. It can also be used as a stand-alone tool to generate animation data. FacEMOTE is our essential next step toward building a consistently expressive body model. Further study will be done in integrating the system onto the body and validating the usefulness of system in the context of consistency.

7 Future Work

This system can be extended to head pose, eye movements, or speech for modeling of a consistently expressive body. The current FacEMOTE system includes head and eye movements within the framework for soft tissue deformation and jaw movements, but they should be modeled separately for proper behavior. We believe head and eye movements are organized with the rest of facial movements such that they are synchronized but are placed apart in the movement hierarchy. We will automate the coefficient and weight table generation process using the current tables as a guide.

The current work opens up a new area of consistently expressive body modeling. FacEMOTE must be integrated with the EMOTE system for gestures. Then the combined system can be properly evaluated to see whether it generates facial expressions consistent with the context of body movements. Another possible area of research is to study whether there is a correlation between the "expressiveness" of facial animation output and generally accepted notions such as emotions and personalities. Machine learning methods should be adopted to build the requisite mappings.

8 Acknowledgements

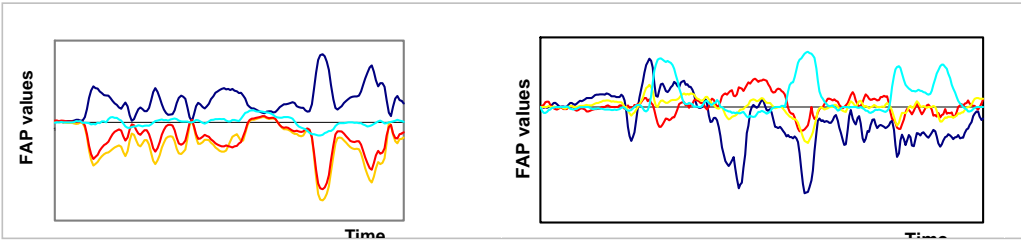
Special thanks go to Catherine Pelachaud for loaning us the MPEG-4 compatible facial model. We acknowledge the MPEG-4 consortium for making MPEG-4 FAP test data available. This

research is partially supported by National Science Foundation IIS99-00297, Office of Naval Research K-5-55043/3916-1552793, and NASA 00-HEDS-01-052. Jan Allbeck, Koji Ashida, and Sooha Lee provided valuable comments and video editing.

References

- BARTENIEFF, I. AND DAVIS, M. A., 1972. Effort-Shape analysis of movement: The unity of expression and function, *Research approaches to movement and personality (Body Movement: Perspectives in Research)*, M. Davis, Ed., Arno Press.
- BARTENIEFF, I. AND LEWIS, D. 1980. *Body Movement: Coping with the Environment*, Gordon and Breach Science Publishers, Inc.
- BLANZ, V. AND VETTER, T. 2000. A morphable model for the synthesis of 3D faces, In *Proceedings of ACM SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 187-194.
- CHI, D., COSTA, M., ZHAO, L., AND BADLER, N. I. 2000. The EMOTE model for Effort and Shape, In *Proceedings of ACM SIGGRAPH 2000*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 173-182.
- DELL, C. 1977. *A Primer for Movement Description using Effort-Shape and Supplementary Concepts*, Dance Notation Bureau Press.
- EKMAN, P. AND FRIESEN, W. V. 1977. *Manual for the Facial Action Coding System*, Consulting Psychologists Press.
- ESSA, I. AND PENTLAND, A. 1997. Coding, analysis, interpretation, and recognition of facial expressions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(7):757-763.
- FAIGIN, G. 1990. *The Artist's Complete Guide to Facial Expression*, Watson-Guptill Publications.
- FRIDA, N. H. AND TCHERKASSOF, A. 1997. Facial action as action readiness, *The Psychology of Facial Expression*, Russel, J. A. and Fernandez-Dols, J. M. Eds., Cambridge University Press.
- ISBISTER, K. AND NASS, C. 2000. Consistency of personality in interactive characters: Verbal cues, non-verbal cues, and user characteristics, *Int. Journal of Human-Computer Studies*.
- KALRA, P., MANGILI, A., MAGNANAT-THALMANN, N., AND THALMANN, D. 1992. Simulation of facial muscle actions based on rational free from deformations, *Eurographics*, 59-69.
- KNAPP M. L. Ed., 1992. *Nonverbal Communication in Human Interaction*, 3rd Ed., Holt, Rinehart and Winston, Inc.
- LASSETER, J. 1987. Principles of traditional animation applied to 3D computer animation, In *Proceedings of ACM SIGGRAPH 1987*, ACM Press / ACM SIGGRAPH, New York, Computer Graphics Proceedings, Annual Conference Series, ACM, 35-44.
- LEE, W. AND MAGNANAT-THALMANN, N. 2000. Fast head modeling for animation, *Eurographics*.
- LEE, Y., TERZOPOULOS, D., AND WATERS, K. 1995. Realistic modeling for facial animation, In *Proceedings of ACM SIGGRAPH 1995*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 55-62.
- HTTP://WWW.LIPSINC.COM
- MPEG-4 SNHC. 1998. Information technology-generic coding of audio-visual objects part 2: Visual, ISO/IEC 14996-2, Final draft of international standard, ISO/IEC JTC1/SC29/WG11 N2501.
- NOH, J. AND NEUMANN, U. 2001. Expression cloning, In *Proceedings of ACM SIGGRAPH 2001*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 277-286.
- OSTERMANN, J. 1998. Animation of synthetic faces in MPEG-4, *Computer Animation*, 49-55.

- PARKE, F. I. 1974. *A parametric model for human faces*, PhD Thesis, University of Utah.
- PARKE, F. I. AND WATERS, K. 1996. *Computer Facial Animation*. A. K. Peters.
- PECK, S. R. 1990. *Atlas of facial expression: An account of facial expression for artists, actors, and writers*, Oxford University Press.
- [HTTP://WWW.MRL.NYU.EDU/~PERLIN/FACEDemo/](http://www.mrl.nyu.edu/~perlin/faceDemo/)
- PIGHIN, F., AUSLANDER, J., LISCHINSKI, D., SALESIN, D. H., AND SZELISKI, R. 1998. Synthesizing realistic facial expressions from photographs, In *Proceedings of ACM SIGGRAPH 1998*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 75-84.
- PLATT, S. M., AND BADLER, N. I. 1981. Animating facial expressions, In *Proceedings of ACM SIGGRAPH 1981*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 245-252.
- TERZOPOULOS, D., AND WATERS, K. 1991. Techniques for realistic facial modeling and animation. *Computer Animation*. Magnanat-Thalmann N., and Thalmann, D. Eds., Springer-Verlag. Ch. 5. 59-73.
- THOMAS, F., AND JOHNSTON, O. 1981. *The Illusion of Life*, Abbeyville Press.
- WATERS, K. 1987. A muscle model for animating three-dimensional facial expression, In *Proceedings of ACM SIGGRAPH 1987*, ACM Press / ACM SIGGRAPH, Computer Graphics Proceedings, Annual Conference Series, ACM, 17-24.
- WEHRLE, T., KAISER, S., SCHMIDT, S., AND SCHERER, K. R. 2000. Studying the dynamics of emotional expression using synthesized facial muscle movements, *Journal of Personality and Social Psychology*, 105-119.
- WILLIAMS, P. L., WARWICK, R., DYSON, M., AND BANNISTER, L. H. 1989. *Gray's Anatomy*, 37th Ed., Churchill Livingstone.



Color plate 1a. FAPs on the mouth region (navy: open_jaw; blue: raise_b_midlip; red: raise_b_lm; yellow: stretch_l_cornerlip)

Color plate 1b. FAPs on 3 regions (navy: raise_l_cornerlip; blue: raise_l_o_eyebrow; red: puff_l_cheek; yellow: lift_l_cheek)



Color plate 2. (a) Original; (b) Indirect+Sustained; (c) Direct+Strong; (d) Free